USEFUL THINGS TO KNOW ABOUT PH.D. THESIS RESEARCH
H. T. KUNG

(Prepared for "What is Research" Immigration Course,
Computer Science Department, Carnegie Mellon University, 14 October 1987)

Presntation Outline

1. Introduction
2. Why Ph.D. thesis could be really difficult for a student
3. Types of Ph.D. theses (from Allen Newell)--not a topic of this talk
4. Growth of a star (the transformation process that some students go through to become a mature researcher)--which stage are you in?
5. Stages of Ph.D. thesis research
6. Methods to get into the depth of a topic (or how to come up with good ideas)
7. Breaking myths
8. Pitfalls to avoid (easy ones to avoid listed first)
9. Some other general advice
10. All the effort is worth it (believe it or not)

1. Introduction

- Ph.D. thesis is treated very seriously at leading universities.

    * Expectation is high.

        - Ph.D. thesis represents a substantial work. Faculty often tell other people that "We have a student working on this area for his or her Ph.D. thesis." Amazingly enough, this is usually sufficient to convince people that the problem is somehow going to be solved.

    * Ph.D. thesis research is a task to ensure that the student can later take on independent, long-term research commitments. (If a Ph.D. student does not intend to be a researcher, the Ph.D. thesis work is not worth the effort in general at least at CMU.)

    * Through the Ph.D. thesis process the student is transformed into a professional researcher.

    * Faculty are judged by the theses of their Ph.D. students.

    * High standard Ph.D. thesis is probably one of the most important factors that contribute to the success of graduate education at leading American universities.

    * Ph.D. thesis is probably the only real challenge for getting a Ph.D. degree.

        - Ph.D. qualifier is seldom a problem for motivated students.

- Ph.D. thesis research is probably more mechanical than a new graduate student would think. (Of course the process is still too complex to be automated.)

* Knowing this mechanism can be more important than thesis results themselves.

* Some information presented here may be relevant to your whole research career, i.e., it is not just for the Ph.D. thesis per se.

- This talk consists of pragmatic advice.

* The talk is based on my personal experience (i.e., not based on any serious research)

- I happen to have research experience in both theory and system areas. We will compare thesis research in these two areas.

* This is a common sense talk and will have down to earth discussions.

- "I wish someone told me this before."

2. Why Ph.D. thesis could be really difficult for a student

- Most likely this is your first, major research experience.

* A big challenge for most students

- No simple recipe

* Different talents

* Different kinds of theses

* Different approaches

- The work is judged by thesis committee (mostly advisor). This produces anxiety.

* Unlike other research you will do, the evaluation mechanism for thesis research is very unique.

* No clear contract

* No clear standard (we only know it is high)

* Recall the Stanford murder case (the former student said, after he had finished--he did finish something-- his jail term, that he might do it again under a similar circumstance).

3. Types of Ph.D. theses (from Allen Newell)--not a topic of this talk

- Opens up new area

- Provides unifying framework

- Resolves long-standing question

- Thoroughly explores an area

- Contradicts existing knowledge

- Experimentally validates theory

- Produces an ambitious system

- Provides empirical data

- Derives superior algorithms

- Develops new methodology

- Develops a new tool

- Produces a negative result

4. Growth of a star (the transformation process that some students go through to become a mature researcher)--which stage are you in?

   - Knowing everything stage

     * Student: "I have designed a supercomputer even before graduate school."

     * Faculty: speechless

   - Totally beaten up stage

     * Student: speechless

     * Faculty: smiling at the student's progress so communication is possible now.

   - Confidence buildup stage

     * Student: "I am not stupid after all." (student thinks)

     * Faculty: "Uh oh, she is ready to argue." (faculty think)

   - Calling the shot stage

     * Faculty: "I am going to design an n-processor supercomputer."

     * Student: "You are crazy, because ..."

5. Stages of Ph.D. thesis research

   a. Selection of area--not a topic of this talk

   b. Selection of advisor--not a topic of this talk

   c. Becoming a researcher in the area

       - Building up general knowledge, experience, and confidence

       - Knowing issues and important questions in the area

       - Capturing research opportunities

         * Don't let any idea or question go by without first giving it careful thought.

             - Be alert and diligent.

         * Pay attention to new technologies

- Examples

    * VLSI, networking, and new chips such as the Weitek floating-point chips three years ago which in some sense gave the initial motivation for the Warp project

- Some useful things to do (from Dave Gifford, MIT)

    * Read recent proceedings of the best conferences, and ask more senior people what were the best papers. Try to figure out what makes a great paper (and thus what makes great research).

    * Keep a notebook that contains your research notes. Put all of your empirical data and initial ideas in the notebook. Make notes on a paper as you read it and think about the assumptions of the author and the importance of the results.

    * Follow references from one paper to another until you know an area extremely well. Don't count on your advisor to hand you all of the relevant papers out of his file drawer. He doesn't have them all!

d. Thesis proposal

- It is the most crucial stage in the sense that the basic concept is worked out here.

    * To get important results you need to ask important questions

    * This is the time you need your advisor most.

    * Problems in later stages are usually rooted from a weak thesis proposal.

- Purpose

    * A research plan

        - A serious attempt to get an overview of the whole research course

        - Not really a contract

            * Need some flexibility because research always has uncertainty.

    * Forming the committee

        - Varies a lot

        - Choose people for your thesis committee that can help with needed expertise. For example, it is useful to have a relevant theory faculty member on a systems committee and vice-versa.

        - However, there is usually no need to optimize too much on the selection of the committee members--advisor still plays the most important role.

- However it can be very important, when

  * you have a "questionable" advisor, or

  * you have an interdisciplinary topic.

* A review

  - If there is any serious doubt, it had better show up now.

  - Proposal could sometimes be viewed as just a forcing function for taking care of certain things.

- Some of the difficult questions always asked in a thesis proposal:

  * What is your approach and what is new?

  * What is your secret weapon? (Herbert Simon)

  * How do you measure your own progress?

  * What are the success or completion criteria?

  * How will the expected results change the-state-of-the-art?

- The grand challenge for a thesis proposal is to come up with an approach or an experiment.

  * It is easy to identify a general problem area, but setting up an approach and designing an experiment can be difficult.

    - Need ideas

      * Just need one good idea, really

      * Unfortunately, there is no magic here (however see some hints below). This is the hard part of any research project for everyone (not just for students).

  * Need independent thinking

    - You should be good enough to start arguing with your advisor on technical issues and research tastes.

  * Need to elaborate on focus, approach, experiment, and potential impact

    - For theory research you may propose some new models of computation.

      * Examples: area-time complexity (new VLSI model in theory), parallel algorithms (new cost models)

    - For system research you may design experiments and argue their relevance.

      * Examples: multiprocessor architecture, compiler for a parallel machine

- Useful things to know when preparing a thesis proposal

    * Be honest. There is no need to exaggerate your claims! If you point out the weaknesses in your approach you will disarm your critics.

    * Pick a project that is manageable so you can do an excellent job - things are always harder than they seem. It is far better to do an outstanding job on a moderate size project than a moderate job on a large project.

    * Include a tentative thesis outline and a month by month schedule in your thesis proposal.

        - This may be difficult to do but it is better than no plan at all.

        - This will also help gauge the total size of the work you are committing yourself to do.

e. Producing results

- Lots of work--what else do you expect?

    * System--be inside an active project without losing sight of thesis

        - Need to be a worker as well as a conceptual person.

        - Your work depends on other people's work and vice versa

            * Opportunity to see real problems

            * Getting good support, including encouragement and demand, from the group

                - It seems that this arrangement really works in all cases.

        - Be quick, because you don't want to be overtaken by the environment (this is one of the pitfalls to avoid, as described below)

    * Theory--be lucky!

        - Be flexible

            * It is hard to insist that you will prove a theorem before you go to sleep.

        - Be quick, because theoretical results are totally portable and so competition can be keen.

- Keep the committee informed (at least those "trouble makers")

    * You can get real help sometimes.

* Committee members are obliged to talk to you.

    - Sometimes finding a qualified person beyond your advisor to discuss your work can be difficult.

* Don't want surprises in the later stage of the thesis

- Ways to finish a thesis

    * Incremental and adaptive approach

        - A sequence of incremental results

    * Big-bang approach (this is not recommended in general)

        - One big theorem

        - A big piece of software or hardware

f. Writing

    - Why some students find that Ph.D. thesis writing is very difficult

        * First major document

        * Writing is time-consuming--part of the .9999 perspiration (Satya)
            - Think how many good sentences you can write in an hour.

            - Fighting with fonts, figures, references, etc.?

                * Please don't be too picky.

        * When results are not totally solid, writing can be really difficult even for an experienced writer (now you know another reason why proposal writing is not easy)

            - Can't say too much and don't want to say any less

            - Writing about flaky results can be a real challenge.

                * In this case you should improve your results first.

        * Writing has to do with presentation rather than finding new results. So writing may not be as exciting..

    - However, thesis writing is useful in the sense that it helps reveal possible problem areas and provides new insights.

        * Help get a large picture on what you really have.

        * Help organize the concepts

* Completeness is forced.

    - You must take care of things that you have been ignoring.

        * For example, you need to do comparison with other results

* Correctness of the results is checked.

    - You had better have the proof now for any plausible "theorem" that you have been believing.

* New insights on how things really work

    - New ways of looking at you results

- Recommendations

    * Get some practice--write some papers before thesis

        - Write some joint papers with people who have substantial writing experience

    * Need to know the theme of the thesis very well

        - Outline first

        - Write the conclusion first (try it at least)

        - Start writing chapters which are more settled.

        - Write the introduction last

        - Iterative process

    * Make the writing as precise as possible, so that you know exactly what you are talking about. This will save lots of rewriting.

        - Precise writing usually also yields good English.

g. Getting final comments from the committee

    - Not too early or too late

        * Getting some committee members to read can be a challenge.

            - They are busy people. You want to give them an "optimal" version to make comments.

    - How much to ask for comments varies a lot

    - Should not have any surprises now.

        * You had better know what you have been doing by now.

        * However, if there is any problem, it had better show up now.

h. Defense

- Mostly a formality and a happy occasion (should be like that)

    * You know that your results are good and you will present them well.

        - You should know the answer to the question – "What are the three main ideas in your thesis?".   You should be able to rattle them off and relate them to previous work.

    * Getting a date set can be more difficult than you think.

        - Committee members do not necessarily stay at CMU as long as you do!

        - Weekend defense is not really desirable.

            * May be difficult to get audience.

- However defense is still very important:

    * Opportunity for final improvements for the thesis

    * Formal presentation to the community

        - Many people form their opinion of your n-years' work from this presentation

    * Presentation material can be used for future presentations

        - Used in recruiting presentations if you have not settled on a job yet

    * Psychologically important

        - Once in a life time occasion--you will remember it always.

    * Don't want to blow it.

        - Absolutely no surprises

i. After defense

- Usually there is still some minor work to be done for the thesis (too bad)

    * Defense was moved early for various reasons

    * New comments from defense

    * Did not have time or did not want to polish the thesis before defense

- Publication

    * Articles, books (or give the thesis to your parents)

    * Very important to publish the results in journals

- This is the only reliable way to archive your
  results.   (You don't want to lose them after
  all these efforts, do you?)

- Publication is important for academic career.

- May break the thesis up in several articles.
  When appropriate, some articles may have joint
  authors such as your advisor.

- Do it right away before you get on to the next
  thing.

* Books can be good too.

- Follow-on work

* Keep mining the thesis--why not?

- Finally you are free!

6. "Methods" to get into the depth of a topic (or how to come up with
   good ideas)

- No magic, but we will still try ....

- How to develop initial ideas

* Study other work and do comparison

- What are similar issues and solutions?

* Look at examples

- Generalization and abstraction

* Make hypothesis and validate it formally or informally--
  keep trying
  - You will discover issues at least.

* Do modeling and abstracting

- Get the essence

* Just do something--be active

- Implementation--details reveal issues

* Join a project to do some real work!

* Handle a smaller case

* Implement a throw-away simulator, language,
  design, etc.

- Start proving "theorems", even if they are known to
  be difficult.

* Quick way to understand issues

* Work with good, experienced researchers (don't forget to
  use your advisor!)

- They might have deep insights on similar problems.

- They can help calibrate the difficulty of the problem.

- You learn the subject matter from them more quickly and directly.

- You learn their techniques

    * Every successful researcher has his or her own bag of "tools":

        - Calculation, synthesis, analysis, persistence

- If they also get stuck once in a while, you know that you are not that bad after all.

- How to develop existing ideas further

    * Exploring problem and solution spaces

        - Enumerate parameters individually (and do quick pruning)

            * To see where your current ideas sit in the space

        - Correlate results

        - Generalize ideas and results to other points in the space

        - Produce phenomena and explain them (Herb Simon)

    * Brainstorming your ideas with others

    * Presenting your ideas in papers or/and seminars

        - Ideas will be checked out carefully and systematically (see above on thesis writing)

    * Example steps that can be used to get some depth from a simple result such as a speed-up curve

        - Explain the curve

        - Look at the problem and solutions spaces

        - Do some comparisons

        - Change the assumptions

            * How stable is the result?

            * How will results vary or correlate under different assumptions?

        - Derive some general principle

            * Similar curves for other situations?

- General comments

* Thinking is the key

    - Thinking is more important than reading

        * Books are not always right.

            - Note that in the system area with few
              exceptions people who build systems do not
              have time nor need to write up their
              experience--it is too bad but it is a
              reality.

    - Be alert on all sorts of opportunities

    - Do the thinking right away while you have it.

        * Ideas and interest may be lost more quickly than
          you like to believe

* Talking to people

    - Don't over do it (you still need to do the work
      yourself)

7. Breaking myths

    a. "Advisor is a stronger researcher than you."

        - It is true that advisor is experienced, wise, smart
          (maybe), and knowledgeable in general. Advisor also sees
          a bigger picture, and has contacts in the area.

        - However, advisor is not always right.

            * Advisor is not as focussed as you.

            * Advisor does not have more time or energy than you
              do.

            * Advisor is not as innovative in general.

                - They know too much.

                - They are more conservative.

                    * They know too many horror stories.

                - Aging does not help.

            * Advisor's knowledge may be obsolete (don't say this
              in front of him or her!).

        - You must believe that you can do better than advisor for
          some research areas.

    b. "System theses take longer than theory theses."

        - The most difficult part of a thesis is to come up with
          some good, new ideas. The difficulty in getting new
          ideas is the same for theory or system research.

            * Theory thesis is in general not about solving open
              problems.

- Actually good theoreticians always work on new problems, models and methods so that they can solve the problems that are "solvable" in the first place.

    * Greatest contributions are ground breaking ones, such as new models.

    * New approaches give new insights to old problems. This is the way open problems usually get solved (e.g., the four-color problem).

* For systems theses it is important that the major ideas in the thesis are independent of the implementation--the goal is to have the ideas live on in other systems as well. A good systems thesis usually has a new algorithm or new method at its core.

* Few theory students who finish really early are likely those who have prior research experience. (Recall that theory results are highly portable!)

* Incompetent theory students are more noticeable than weak system students. So we don't often see theory students who drag on for a long time.

- There are some differences in systems and theory research however, but they should not have too much impact on the thesis research time.

    * System needs implementation, whereas theory needs more background study.

    * Theory research is self-sufficient and system implementation may depend on other people's work (you should not get into a situation where you don't have control).

c. "Ph.D. thesis research follows some standard guidelines."

- Yes, a Ph.D. this must represent a substantial result in a very high standard.

- But there are many ways to leave a mark in a research area. As long as you have come up with some good ideas and pushed the frontier of knowledge, you will be surprised sometimes how flexible your committee could be in terms of the research approach, acceptable results, and thesis presentation.

- There is a small percentage of Ph.D. theses completed in unusual manner. Don't give up too early if you belong to this class. Try it or you will never know.

8. Pitfalls to avoid (easy ones to avoid listed first)

a. The goal is too big to reach.

- Theory

    * Proving P /= NP

                * Proving P = NP is even worse (likely this thesis
                  will never finish!).

                * Deciding whether P = or /= NP is best of the three
                  (i.e., be flexible)

       - System

            * The initial effort is so large that real issues
              never get a chance to be looked at.

            * It is important to size the project and evaluate the
              total effort carefully based on past experiences.

b. Ideas cannot stand without an implementation that competes
   with commercial products.

      - Chess machine implementation is OK, because there is no
        commercial competitor.

      - In this sense, Warp hardware is more difficult than
        software.

      - Floating-point designs that require a high-performance
        chip implementation to validate the concept would be
        disastrous.

      - Never need to implement another vector processor!

c. The thesis area is overtaken by technology and environment

      - Technology advances have solved the thesis problem.

          * A clever operating system using no more than 128K
            memory is not very interesting today.

      - Advisor (or student sometimes) has changed his or her
        interest

      - Other new projects have better approaches and
        opportunities

      - Other people have published similar and/or better
        results.

      - Advisor has a better job elsewhere or the project is
        over.

      - Lesson: You should always do your thesis as quickly as
        possible.

d. Totally isolated work

      - No encouragement and support--no one cares about your
        thesis

          * Can't even find an advisor sometimes

          * Doing a thesis away from CMU is really difficult.

      - System research

          * Lone ranger approach is almost suicidal.

- No software, systems and application support for evaluation

- Very difficult to do anything real without feedback from a community

- Theory research

  * At least global networking is needed.

e. Not knowing when to stop

- Thesis is not the last research you will do.

- You can do the same research after your Ph.D. thesis (while making more money).

- Learn to make reasonable assumptions to restrict the problem

f. Unhealthy competition between student and advisor

- This is more likely to happen in the theory area.

- The potential is always there (especially for smart professors with lots of ego). In general if both sides try to be fair, things can always be worked out.

g. Lots of numbers and hacking but no fundamental principles

- System research has to have more than implementation.

- Implementation for a thesis research is interesting only if it can be used to validate some theory.

- This problem should be fixed as early as possible.

h. Things dragged on--wonderful general ideas in the beginning that never get developed into a coherent approach (i.e., heading to a black hole--there is no output)

- Wrong areas for the student (and perhaps the advisor) with respect to ability and interest

- Nightmare case--it does no good to anyone.

9. Some other general advice

- Stay away from areas that have been thoroughly mined by your ancestors.

  * Keep yourself at the very front of a research area so that you have a better chance to hit something big or at least new.

  * After all in research what matters is the work that pushes us into new territories.

  * Make use new advances in other areas

- Don't avoid thinking

  * Thinking is hard but there is no substitute for it.

- Psych yourself up for this unique experience of doing a Ph.D. thesis

    * Make yourself believe you are solving the most important problem in the world

    * Remember what worked for you before

        - If you work best when you are competing with others, then create some confrontation.

    * Must be very alert about issues and opportunities

    * Thesis process is sort of artificial (almost a torture in some way)

        - The thesis is judged by a committee (mainly your advisor)

            * More subjective than exams

        - Probably one of the most humiliating experiences for people of this age (advisors should all remember this and be considerate.)

        - The process is not a typical research style--you don't do anything similar to it again even if you will be doing research after the degree.

    * The thesis process can be long and treacherous. (Be prepared for it.)

        - You don't want depression.

    * There are quite a few very competent people who just do not want to go through this.

- Use forcing functions well to speed up the thesis process

    * Competing with someone else

    * Family pressure

    * Financial pressure

    * A job is waiting

    * Advisor is leaving or project is over

    * Equipment is retiring

- Never throw away advisor's comments

    * Cox-Denning case

- Keep good relationship with your advisor (even after you graduate)

    * Good thing to do--no exception almost

    * Relationship is unique.

        - Advisor usually has lots of influence on you in this

very important stage of your life. Advisor also appreciates the good research you did with him, and is in general interested in your well-being.

* Advisor may be your mentor for your entire career.

10. All the effort is worth it (believe it or not)

- Experience from Ph.D. thesis research is unique. You have learned how to do research. Future research is going to be more interesting because you will know how to do it, so you will have more freedom and fun.

- Almost all leaders in research have this experience. You will have confidence in your research ability. You will look at things differently than people who did not go through this process. It is very clear that Ph.D. thesis research is still the best way we know of in developing powerful researchers.

- In summary, it is the best investment for becoming a successful researcher.